Computational Algorithms and Numerical Dimensions

www.journal-cand.com

Com. Alg. Num. Dim.Vol. 4, No. 1 (2025) 34–47.

Paper Type: Original Article

Advancing Natural Language Processing for Persian Movie Review Analysis: Roadmap and Opportunities

Ramin Safa[1,*], Sedigheh Kaveh[1]

[1] Department of Computer Engineering, Ayandegan Institute of Higher Education, Tonekabon, Iran; r.safa@outlook.com; kaveh@morvarid-systems.or.

Citation:

Received: 11 December 2024 *Revised: 13 Februray 2025* *Accepted: 28 Februray 2025*	Kaveh, S., & Safa, R. (2025). Advancing natural language processing for Persian movie review analysis: roadmap and opportunities. *Computational algorithms and numerical dimensions*, *4*(1), 34-47.

Abstract

As Persian-language movie platforms gain popularity, analyzing user-generated content becomes increasingly important. Advanced Natural Language Processing (NLP) tools, such as TextBlob, NLTK, VADER, Num2faword, PersianTools, Parsivar, Hazm, and BERT, provide robust methods for sentiment analysis, text preprocessing, and aspect-based analysis tailored to Persian movie reviews. These tools address unique challenges, including diverse writing styles, non-standardized sentiment lexicons, and linguistic complexities of Persian. This paper presents a roadmap for developing NLP-based solutions, highlighting these tools' applications with example outputs. The integration of these tools into sentiment analysis pipelines offers significant opportunities for improved user experiences, personalized recommendations, and actionable insights for platform owners. By addressing challenges and capitalizing on the potential of advanced NLP techniques, this research aims to foster the growth of Persian-language movie platforms and contribute to their global competitiveness.

Keywords: Sentiment analysis, Natural language processing, Persian language processing, Movie review analysis, Recommender systems.

1 | Introduction

The rapid growth of Persian-language movie and streaming platforms has led to a significant increase in the volume of user-generated content, such as reviews and comments, across these platforms [1]. Platforms like Filimo[1], Namava[2], Filmnet[3], Aparat[4], Rubika[5], and Telewebion[6] are popular examples that actively enable

[1] https://filimo.com
[2] https://namava.ir
[3] https://filmnet.ir/
[4] https://aparat.com/
[5] https://web.rubika.ir/
[6] https://telewebion.com/

Corresponding Author: r.safa@outlook.com
https://doi.org/10.22105/cand.2024.493386.1168

users to share their thoughts and feedback through comments and reviews. As the volume of this Persian-language content continues to expand at an unprecedented rate, the need for effective Natural Language Processing (NLP) techniques to analyze and extract valuable insights from this data has become increasingly crucial. This user-generated content represents a rich and largely untapped resource that can provide invaluable feedback for platform improvements, personalized recommendations, and a deeper understanding of viewer preferences and sentiments [2].

However, the existing body of research on sentiment analysis, recommender systems, and other NLP-powered tools for Persian movie reviews remains severely limited [1]. While substantial progress has been made in applying these techniques to user-generated content in other languages [3], the unique linguistic and cultural nuances of the Persian language have presented distinct challenges that have yet to be thoroughly addressed. Overcoming these challenges and developing robust NLP capabilities for Persian movie reviews is not only a significant academic and technological challenge but also a crucial step in ensuring the continued growth and competitiveness of the Persian-language movie industry.

This paper aims to outline a comprehensive roadmap for developing advanced NLP-based solutions to analyze Persian movie reviews and enable more effective user recommendation and platform improvement. We first provide an in-depth review of the state-of-the-art in NLP techniques that have been applied to user-generated content analysis in other languages, highlighting their potential applicability to the Persian movie review domain. By understanding the progress made in related fields, we can better identify the necessary adaptations and innovations required to address the specific needs of the Persian movie review landscape. Next, we investigate the specific challenges of working with Persian-language movie reviews, such as handling diverse writing styles, lack of standardized sentiment lexicons, and adapting techniques designed for other languages. We also discuss the significant benefits that effective NLP-powered tools can bring to Persian movie platforms, including enhanced user experience, personalized recommendations, and valuable insights for platform owners. Based on this analysis, we propose a step-by-step roadmap for future research and the development of novel Persian-specific models and architectures. By implementing this roadmap, researchers and practitioners can drive the advancement of NLP capabilities in the domain of Persian movie review analysis.

2 | Related Work

Sentiment analysis has become a prominent area of research, fueled by the abundance of data available online. Popular techniques include Naïve Bayes classifiers, Support Vector Machines (SVMs), Latent Dirichlet Allocation (LDA), and neural networks, often preceded by preprocessing steps such as tokenization, Part-of-Speech (POS) tagging, and parsing. These methods are applied to various levels of opinion mining, including document and sentence-level analysis. Challenges arise when reviews contain mixed sentiments, making it difficult to classify their overall tone. Early studies, like Peng et al. [4], demonstrated that statistically derived keyword lists outperform human-curated lists in sentiment classification. Advances in NLP also extend to applications like social media data mining and medical text analysis, where models such as Hidden Markov Models (HMMs) and Conditional Random Fields (CRFs) are commonly employed.

NLP methods have been extensively applied to analyze movie reviews, with classifiers like SVMs, random forests, and logistic regression achieving high accuracies. For instance, hybrid models, such as naïve Bayes combined with Genetic Algorithms (GA), have attained over 93% accuracy. Researchers have also used feature extraction techniques with unigrams, bigrams, and trigrams to enhance classification performance. Neural networks, including Recursive Neural Tensor Networks (RNTNs), have further improved sentiment analysis but often require significant computational resources. To address this, models like decision trees, random forests, and SVMs are frequently used due to their efficiency and robust performance. These methods strike a balance between computational feasibility and accuracy, making them suitable for many sentiment analysis tasks.

While advancements in NLP for user-generated content analysis in other languages have been made, limited research has been conducted on Persian-language movie reviews. *Table 1* compares the overall NLP research on user-generated content analysis across different languages based on a meta-analysis [5].

Table 1. Comparison of NLP research for user-generated content analysis across different languages.

Metric	Persian	English	Chinese	Arabic
Sentiment analysis	Limited	Advanced	Moderate	Emerging
Topic modeling	Sparse	Extensive	Moderate	Emerging
Recommender systems	Minimal	Mature	Moderate	Limited
Dataset availability	Scarce	Abundant	Growing	Increasing
Linguistic complexity	High	Moderate	Moderate	High

As shown in the table, the research on NLP techniques for Persian-language user reviews, including movie reviews, is relatively underdeveloped compared to other major languages. The limited availability of high-quality Persian language datasets, coupled with the inherent linguistic complexities of the Persian language, has delayed progress in this area. *Table 2* provides an overview of the commonly used methods and tools in sentiment analysis. This table highlights the potential approaches that could be adapted and applied to the analysis of Persian-language movie reviews.

Table 2. Methods and tools in sentiment analysis.

Method	Description	Tools
Lexicon-based	Use pre-compiled sentiment lexicons to determine sentiment polarity	VADER [6], SentiWordNet [7], TextBlob [8]
Machine learning	Train supervised models for sentiment classification	Logistic regression [9], SVM [10], LSTM [11], BERT [12]
Hybrid	Combine lexicon-based and machine-learning approaches	SenticNet [13]

By leveraging the lessons learned from the progress made in other languages and tailoring the techniques to the unique characteristics of the Persian language, we can catalyze the development of NLP-powered solutions for the Persian movie review ecosystem.

3|Text Classification for Persian Movie Platform Sentiment Analysis

The booming Persian movie industry and the rise of online movie platforms have led to a flood of user-generated content, including reviews, ratings, and discussions. Analyzing this textual data can provide valuable insights into user sentiment and preferences, informing platform improvements, content curation, and marketing strategies.

3.1|Problem Statement and Objectives

The main goal of this research is to create a reliable and accurate sentiment analysis model for a popular Persian movie platform. The platform's data includes the titles of movies and TV series, user reviews and ratings/scores, metadata such as categories, genres, actors, and crew information, as well as user profiles and interaction data (e.g., viewing history, recommendations) [14]. This rich dataset encompasses a wide range of elements, as depicted in *Fig. 1*.

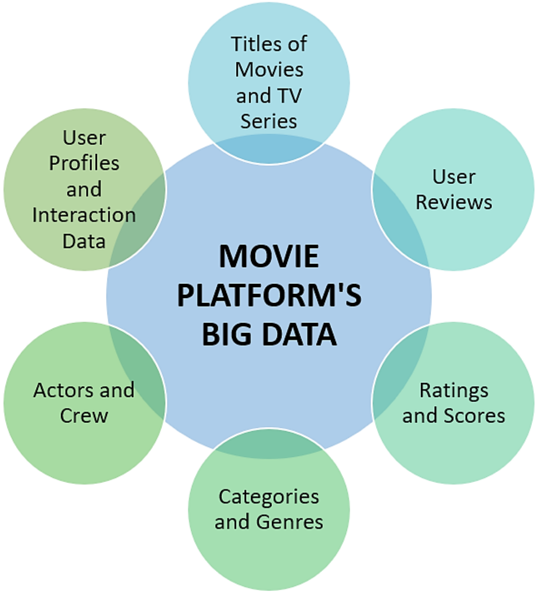

Fig. 1. Persian movie platform big data.

Analyzing this diverse and large-scale dataset presents both challenges and opportunities for sentiment analysis, as the researchers aim to capture the tones of how Persian users express their opinions and emotions about the platform's content.

By developing an effective sentiment analysis model for this Persian movie platform, the researchers can enable a range of applications, such as personalized content recommendations, targeted marketing campaigns, and improved customer experience [15]. Additionally, the insights gained from this study can contribute to the broader understanding of sentiment analysis in the Persian language and inform future research in this domain.

3.2 | Data Acquisition and Preprocessing

To build and train our sentiment analysis model, we need a comprehensive dataset of Persian movie reviews. This dataset will be the foundation for our model's ability to classify sentiments accurately.

For data acquisition, reputable Persian movie platforms and forums can be identified to gather user-generated content. Web scraping techniques or APIs can be employed to extract reviews, ratings, discussions, and other relevant data [16]. The data should be cleaned by removing irrelevant or low-quality content, such as spam or promotional posts. Finally, the cleaned data can be stored in a suitable format, such as CSV or JSON, for further analysis. The data preprocessing process involves several critical steps to prepare textual data for analysis. This ensures that the data is in a suitable format and of high quality, ultimately leading to more accurate insights and results. *Table 3* outlines the key steps in data preprocessing, including text cleaning, tokenization, stop word removal, stemming, feature extraction, and data splitting.

Table 3. Key steps in data preprocessing.

Step	Description
Text cleaning	Remove HTML tags, URLs, and special characters. Handle inconsistencies in text formatting (e.g., spacing, punctuation).
Tokenization	Break down text into words or n-grams (sequences of words). Use Persian-specific tokenizers to handle complex linguistic structures.
Stopword removal	Eliminate common words that do not carry significant semantic meaning (e.g., "و", "که", "از").
Stemming/Lemmatization	Reduce words to their root forms to improve feature representation. Use Persian-specific techniques to handle morphological variations.
Feature extraction	Employ techniques like bag-of-words, TF-IDF, or word embeddings to represent text as numerical features.
Data splitting	Divide the dataset into training and testing sets to evaluate model performance.

Throughout the data preprocessing stage, researchers would carefully monitor the impact of each step on the data quality and the performance of the classification models.

The workflow is depicted in *Fig. 2*, offering a step-by-step roadmap for text processing, spanning from data acquisition to its practical application.

3.3 | Machine Learning Techniques and Tooling

3.3.1 | Traditional machine learning approaches

Traditional machine learning algorithms provide a solid foundation for sentiment analysis. These methods rely on statistical techniques to classify text into sentiment categories.

I. Logistic regression: a statistical model that estimates the probability of a review belonging to a specific sentiment class. It is a simple yet effective model that can be used as a baseline [17].

II. Support Vector Machines: a powerful classification algorithm that finds the optimal hyperplane to separate positive and negative reviews. SVMs are particularly effective when dealing with high-dimensional data [10].

III. Naive bayes: a probabilistic classifier based on Bayes' theorem, assuming feature independence. It is a simple and efficient algorithm, but it may not perform well when features are highly correlated [18].

IV. Decision trees and random forests: ensemble methods that combine multiple decision trees to improve accuracy. Decision trees can be visualized to understand the decision-making process of the model [19].

3.3.2 | Deep learning approaches

Deep learning techniques, particularly neural networks, have revolutionized the field of NLP. For sentiment analysis, the following deep learning architectures can be considered as follows.

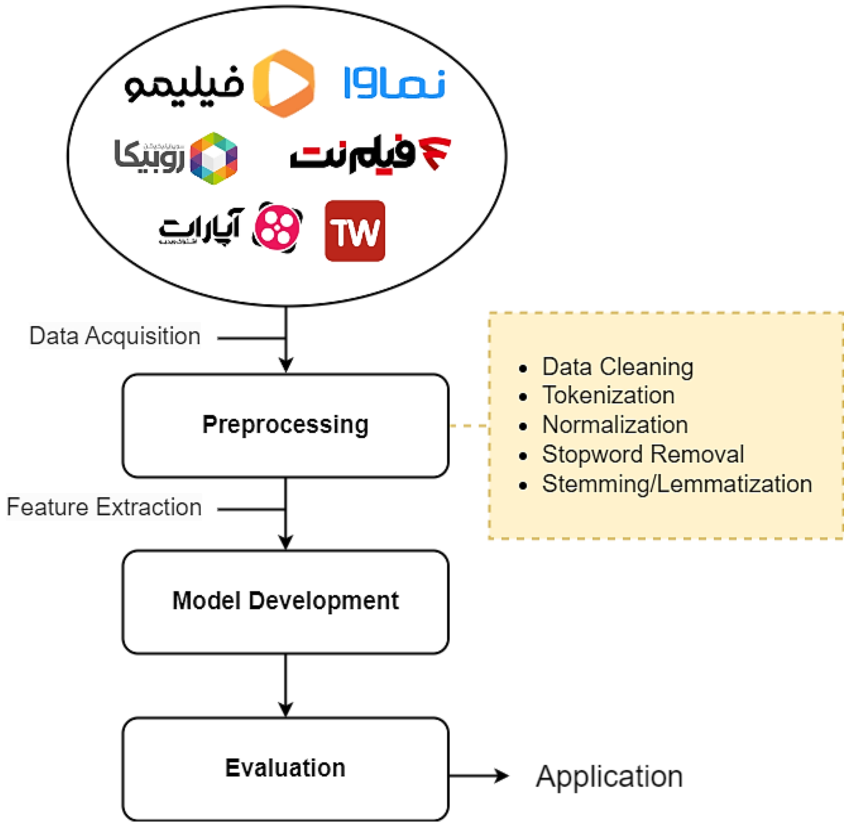

Fig. 2. Roadmap for text processing in Persian movie review analysis.

I. Recurrent Neural Networks (RNNs): RNNs are well-suited for sequential data like text, capturing long-range dependencies. However, they can suffer from vanishing gradient problems, especially for long sequences [20].

II. Long Short-Term Memory (LSTM) networks: a type of RNN that can handle long-term dependencies more effectively than traditional RNNs. LSTMs use memory cells to store information over time, allowing them to capture the context of the text [21].

III. Transformer-based models: transformer-based models, such as BERT, have achieved state-of-the-art performance on various NLP tasks, including sentiment analysis. They utilize self-attention mechanisms to weigh the importance of different parts of the input sequence [22].

3.3.3 | Lexicon-based approaches

Lexicon-based methods rely on sentiment lexicons, which are dictionaries of words and their associated sentiment scores [23]. Tools like TextBlob can be used to perform sentiment analysis based on these lexicons. By combining these techniques and addressing the specific challenges of Persian language sentiment analysis, we can develop a robust and accurate model.

3.4 | Evaluation Methods

To assess the performance of the sentiment analysis models, a range of evaluation metrics and techniques will be employed [24]. Accuracy measures the proportion of correctly classified instances out of the total number of instances. It is calculated as follows.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}).$$

While accuracy provides a general overview of the model's performance, it may not be sufficient in the case of imbalanced datasets.

Precision, recall, and F1-score are more nuanced metrics that consider the specific types of errors made by the model. Precision measures the proportion of true positive instances among the positive predictions, recall

measures the proportion of true positive instances that were correctly identified, and the F1-score is the harmonic mean of precision and recall, providing a balanced measure of the model's performance. They are calculated as follows.

$$\text{Precicion} = \text{TP} / (\text{TP} + \text{FP}),$$
$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}),$$
$$\text{F1} = 2 * (\text{Prec} * \text{Rec}) / (\text{Prec} + \text{Rec}).$$

The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) provide a more comprehensive assessment of the model's ability to distinguish between different sentiment classes [25]. The ROC curve plots the true positive rate (sensitivity) against the false positive rate (specificity) at various threshold settings. The AUC measures the overall performance of a classification model, indicating its ability to distinguish between positive and negative classes.

To ensure the robustness and generalizability of the sentiment analysis models, cross-validation techniques, such as K-fold cross-validation, will be employed [11]. This involves repeatedly training and evaluating the model on different subsets of the data, providing a more reliable estimate of the model's performance than a single train-test split.

Human evaluation can be used to assess the quality and accuracy of the model's predictions. Domain experts and platform users can manually evaluate a sample of the model's predictions and provide feedback on their correctness and relevance.

By employing a comprehensive set of evaluation methods, researchers can gain a thorough understanding of the sentiment analysis model's performance and identify areas for improvement.

4 | Tools and Techniques in Persian Sentiment Analysis

This section outlines the tools and techniques essential for processing and analyzing Persian movie platform comments. These tools address key challenges in Persian-language sentiment analysis, such as tokenization, normalization, aspect detection, and polarity classification. By leveraging these tools, researchers can tackle the unique complexities of the Persian language, including its diverse writing styles and linguistic nuances. Examples of their outputs are provided to illustrate their practical applications and effectiveness.

4.1 | Preprocessing Tools

Preprocessing is a critical step in Persian sentiment analysis to handle linguistic complexities and ensure clean input for downstream tasks [26].

4.1.1 | Parsivar

Parsivar[1] is a comprehensive toolkit designed specifically for preprocessing Persian text. It provides functionality for normalization, tokenization, stemming, and POS tagging [27]. These features address common challenges in Persian text, such as inconsistent character encoding, word boundary errors, and complex morphological variations.

Key features of Parsivar include its ability to standardize text by resolving issues with half-spaces, diacritics, and inconsistent spellings. It also supports space correction, where improperly spaced words are adjusted to improve readability and semantic accuracy. For instance, it can correctly interpret and tokenize "برنامه نویس" as "برنامه‌نویس." Parsivar’s stemming functionality effectively identifies the root forms of both verb and non-verb words, which is critical for sentiment analysis and other NLP tasks. Parsivar’s applications are broad, ranging from improving the accuracy of sentiment analysis to enhancing machine translation and information retrieval systems. For example, by preprocessing raw text, Parsivar ensures clean input for deep learning

[1] https://github.com/ICTRC/Parsivar

models, ultimately improving classification accuracy. *Table 4* provides examples of text preprocessing using Parsivar, including normalization, stemming, and tokenization.

Table 4. Examples of text preprocessing with parsivar.

Feature	Description	Example
Normalization	Standardizes text by resolving issues with half-spaces, diacritics, and inconsistent spellings.	"برنامه نویس" → "برنامه‌نویس"
Stemming	Reduces words to their root form, distinguishing between verbs and non-verbs.	"ستارگان" → "ستاره"
Tokenization	Breaks text into tokens such as words or sentences for better processing.	"این فیلم عالی بود." → ["این", "فیلم", "عالی", "بود"]

4.1.2 | Hazm

Hazm1 is another widely used preprocessing tool for Persian text, offering functionalities similar to Parsivar but with a simpler and more accessible interface. Hazm is particularly suited for projects requiring quick and efficient normalization and tokenization without the need for complex configurations [28].

While Hazm lacks advanced features like dependency parsing, it excels in basic preprocessing tasks such as stemming, removing stopwords, and standardizing characters. For instance, it normalizes text by converting Arabic characters, such as "ي " to Persian equivalents like "ی," and replaces non-standard half-spaces with proper ones. Hazm's simplicity makes it a preferred choice for beginners and lightweight applications. It is particularly useful in educational settings or smaller-scale sentiment analysis projects where complex preprocessing requirements are not a priority.

4.2 | Sentiment Analysis Tools

These tools analyze sentiment using various approaches, including lexicon-based and rule-based methods. *Table 5* provides examples of text preprocessing using Hazm.

Table 5. Examples of text preprocessing with hazm.

Feature	Description	Example
Normalization	Converts Arabic characters to Persian equivalents and handles spacing issues.	"ي " → "ی"
Stemming	Removes affixes to find the root word.	"دویدم" → "دویدن"
Stopword removal	Eliminates common words that do not carry significant meaning.	" این فیلم خیلی خوب بود" → "خوب" , "فیلم"

4.2.1 | TextBlob

TextBlob2 is a widely used, lexicon-based sentiment analysis tool that calculates the polarity and subjectivity of text. Polarity is measured on a scale from -1 (negative) to +1 (positive), while subjectivity is rated between 0 (objective) and 1 (subjective) [8]. TextBlob is easy to use and suitable for analyzing short texts, like movie reviews, and provides quick insights into the sentiment of a piece of text.

Key features of TextBlob include its simplicity and the use of pre-built lexicons to classify text. However, it may struggle with the nuances of the Persian language, especially with idiomatic expressions or culturally specific sentiments. Despite these limitations, it remains a useful tool for initial sentiment analysis and can be a good starting point for processing Persian movie reviews. TextBlob can be used in various applications, such as analyzing user comments to determine customer satisfaction or classifying movie reviews as positive or negative. The examples of sentiment analysis using TextBlob are presented in *Table 6*.

[1] https://github.com/nasim-fani/hazm

[2] https://github.com/sloria/TextBlob

Table 6. Examples of sentiment analysis with TextBlob.

Comment	Translation	Polarity	Sentiment
"فیلیمو باعث شده من از تلویزیون‌های سنتی فاصله بگیرم".	"The variety of genres in Filimo has kept me away from traditional TV."	0.0	Neutral
"دوبله اختصاصی فیلیمو بسیار با کیفیت و حرفه‌ای است".	"Filimo's exclusive dubbing is very high quality and professional."	0.55	Positive
"پشتیبانی فنی ضعیف است و پاسخگو نیست".	"Technical support is weak and unresponsive."	-0.18	Negative

4.2.2 | Valence Aware Dictionary for Sentiment Reasoning (VADER)

VADER1 is a lexicon-based sentiment analysis tool specifically designed to handle the nuances of social media text [29]. It accounts for punctuation, capitalization, and degree modifiers to improve sentiment accuracy in informal texts, making it well-suited for analyzing Persian movie reviews.

VADER assigns a sentiment score that indicates the intensity of sentiment in the text. It also considers context, so words like "not" or "very" can adjust the polarity. For example, VADER can recognize the negative sentiment in a comment like "دسته بندی ژانر دقیق نیست," adjusting the sentiment score accordingly to reflect dissatisfaction.

VADER's ability to adapt to different intensifiers (e.g., "خیلی" ,"بسیار") and negations (e.g., "بد" ,"نه") makes it particularly useful for analyzing Persian-language content, where such linguistic features often change the sentiment of a statement. *Table 7* shows VADER sentiment analysis results, including comment translations, polarity scores, and sentiment (negative, neutral, positive).

Table 7. Sentiment analysis results using VADER.

Comment	Translation	Polarity	Sentiment
"دسته بندی ژانر دقیق نیست و یافتن فیلم‌ها زمان می‌برد".	"Genre categorization is inaccurate, and finding movies takes time."	-0.45	Negative
"آرشیو سریال‌های ایرانی بسیار کامل است".	"The archive of Iranian series is very complete."	0.0	Neutral
"دوبله اختصاصی فیلیمو بسیار با کیفیت و حرفه‌ای می‌باشد".	"Filimo's exclusive dubbing is very high quality and professional."	0.76	Positive

4.3 | Language-Specific Tools

These tools address specific issues in Persian text processing, such as the handling of numbers, dates, and structural elements that require attention to linguistic details.

4.3.1 | Num2faword

Num2faword2is a Python library designed to convert numbers into their Persian word equivalents. This is crucial in many applications, such as processing movie reviews that mention prices, dates, or statistics [28]. By converting numbers to words, Num2faword ensures consistency in how numerical information is presented in Persian texts.

[1] https://github.com/cjhutto/vaderSentiment

[2] https://github.com/5j9/num2fawords

For example, the number "123456" would be converted to "صد و بیست و سه هزار و چهارصد و پنجاه و شش," ensuring that it aligns with the linguistic norms of the Persian language. This tool is especially useful for applications involving financial transactions, subscription models, or any platform where numerical data is a significant part of the analysis.

4.3.2. | PersianTools

PersianTools1 is a collection of functions and utilities designed to address various Persian text-processing challenges [28]. These tools include functionalities like date conversion (e.g., Gregorian to Persian dates), validation of national ID numbers, and handling various formats in Persian texts. These operations are essential in real-world applications, such as movie platforms, where accurate data formats are required.

For example, PersianTools can convert dates in the Gregorian calendar (e.g., 2024-10-31) into the Persian calendar format (e.g., 1403-08-09), ensuring that reviews involving dates or time-sensitive information are handled correctly. This tool also helps to validate and standardize other data, such as phone numbers and ID numbers, making it a comprehensive solution for text processing in Persian-language platforms.

4.4 | Advanced Machine Learning Models

These models utilize deep learning techniques and transformer-based architectures to enhance the accuracy of sentiment analysis, particularly in dealing with the complex structures and contextual meaning of Persian text.

4.4.1 | BERT

BERT2 (Bidirectional encoder representations from transformers) is a powerful deep learning model developed by Google that has revolutionized NLP. Unlike previous models, BERT uses a bidirectional approach to process text, meaning it reads the text in both directions (left-to-right and right-to-left) simultaneously [12]. This approach allows BERT to capture a deeper understanding of word meanings in context, making it especially effective for analyzing Persian movie reviews, where the meaning of words can change depending on the surrounding text.

BERT is pre-trained on large corpora and can be fine-tuned for specific tasks, such as sentiment analysis. For Persian-language content, BERT's ability to handle long-range dependencies and complex syntax makes it a powerful tool for extracting sentiment from nuanced movie reviews. It is trained on the task of predicting masked words in a sentence, allowing it to understand contextual relationships between words better.

For example, BERT might classify the sentiment of the Persian sentence " توانایی تماشای فیلم به صورت آفلاین عالی است" (the ability to watch movies offline is excellent) with a positive sentiment, yielding a polarity score of +0.68. This performance reflects BERT's ability to understand not just individual words but their relationships in the context of the entire sentence. *Table 8* provides examples of sentiment analysis using BERT, showing the translation of comments, their polarity scores, and the corresponding sentiment (positive or neutral).

4.4.2 | Transformer-based models (GPT-3, GPT-4, mT5, etc.)

Recent advancements in transformer-based models have significantly progressed natural language processing. Beyond earlier models like GPT-2[3] and T5[4], newer models such as GPT-3[5], GPT-4[6], mT5 (a multilingual version of T5), and PerT5[7] (a Persian-specific version of T5) offer enhanced capabilities in language understanding and text generation. For example, GPT-3 and GPT-4, with their complex architecture and extensive training datasets, excel in generating high-quality text, translating documents, and performing

[1] https://github.com/topics/persian-nlp
[2] https://github.com/google-research/bert
[3] https://github.com/openai/gpt-2
[4] https://huggingface.co/docs/transformers/en/model_doc/t5
[5] https://github.com/openai/gpt-3
[6] https://github.com/topics/gpt-4
[7] https://github.com/ymcui/PERT/blob/main/README_EN.md

sophisticated text analyses. These models are particularly effective in conversational AI and sentiment analysis. Meanwhile, mT5, designed for multilingual tasks, demonstrates strong performance in summarization and translation, including tasks for the Persian language. Additionally, PerT5 and ParsBERT[1], specifically developed for Persian, address the language's structural challenges, achieving higher accuracy in sentiment analysis and user review processing.

Table 8. Examples of sentiment analysis using BERT.

Comment	Translation	Polarity	Sentiment
"توانایی تماشای فیلم به صورت آفلاین عالی است."	"The ability to watch movies offline is excellent."	0.68	Positive
"پشتیبانی مشتریان فیلیمو گاهی اوقات پاسخگو نیست."	"Filimo's customer support is sometimes unresponsive."	0.51	Neutral

Applications in Persian Sentiment Analysis

These transformer-based models can accurately interpret complex sentiments in Persian texts. For instance, the PerT5 model can analyze a sentence like: "فیلم‌های جدید خیلی تکراری شده‌اند و داستان‌ها جذابیت ندارند" ("the new movies have become repetitive, and the storylines are not engaging") and identify it as expressing slightly negative sentiment. This capability is attributed to their ability to process longer texts and capture semantic nuances, making them highly effective for tasks requiring detailed text comprehension.

Integration with other tools

Advanced transformer-based models can be combined with other methods, such as lexicon-based tools (e.g., VADER or Text Blob), to improve both the speed and accuracy of sentiment analysis. This hybrid approach is especially beneficial for large-scale review analysis systems like Persian movie platforms.

4.5 | Hybrid Approaches

While the tools and models discussed above are effective on their own, hybrid approaches can combine multiple models to enhance the accuracy and robustness of sentiment analysis systems [30]. For example, integrating traditional lexicon-based tools like VADER or TextBlob with advanced transformer models like BERT allows for greater flexibility and accuracy in handling both simple and complex sentiment tasks.

A hybrid system might first use VADER to quickly classify obvious positive or negative reviews based on common keywords, then refine the results using BERT for a more nuanced sentiment analysis of reviews with mixed or subtle opinions. This two-tier approach ensures that the sentiment analysis process remains both efficient and highly accurate, particularly for large-scale platforms like Persian movie streaming services. This hybrid approach can lead to better results in real-world sentiment analysis tasks, where simple, rule-based models might fail to capture the complexity of the Persian language and culture.

5 | Future Directions and Recommendations

To further advance the field of Persian movie review analysis, several key areas require continued research and development.

Dataset development and annotation: identifying and curating high-quality datasets of Persian movie reviews from various online sources (e.g., movie streaming platforms, review websites) is crucial. Annotating these datasets with sentiment labels, aspect categories, and other relevant metadata will enable supervised model training.

Model adaptation and development: evaluating the performance of state-of-the-art sentiment analysis, aspect extraction, and recommendation models developed for other languages on Persian movie reviews will help

[1] https://github.com/hooshvare/parsbert

identify key challenges and limitations [1]. Strategies can then be developed to fine-tune and adapt the models to handle better the unique characteristics of the Persian language and movie review data. Designing novel neural network architectures and training strategies specifically tailored for Persian language processing, incorporating domain knowledge, and exploring the integration of domain-specific features can further enhance the performance of these approaches.

Evaluation and benchmarking: establishing standardized evaluation protocols and benchmark datasets will be crucial to assessing the performance of Persian movie review analysis models. Thorough testing and comparison of the adapted and novel techniques should be conducted to identify the most effective approaches.

Industry collaboration and deployment: engaging with the Persian movie industry to obtain feedback and measure the real-world impact of the developed solutions is essential [31]. Integrating NLP-powered tools into the platforms' infrastructure to provide real-time analysis and insights will facilitate their adoption and use. Continuous monitoring of the performance of deployed models and iterative improvement based on user feedback and new data will be crucial for their long-term success.

By following these directions, researchers and practitioners can contribute to the advancement of NLP for Persian movie review analysis, ultimately enhancing the overall growth and competitiveness of the Persian-language movie industry.

6 | Conclusion

The rapid growth of the Persian-language movie industry has led to a significant increase in user-generated content, such as reviews and comments. Effectively analyzing this data using advanced NLP techniques can provide invaluable insights to platform owners, content creators, and users. However, the existing research on NLP for Persian movie reviews remains limited, presenting both challenges and opportunities.

This paper has outlined a comprehensive roadmap for developing robust NLP-based solutions to address the unique needs of the Persian movie review landscape. By curating high-quality datasets, adapting state-of-the-art techniques, and designing novel Persian-specific approaches, researchers can drive progress in this domain. The successful implementation of this roadmap has the potential to transform the Persian movie ecosystem, empowering platform owners, content creators, and users with enhanced user experience, personalized recommendations, and valuable industry insights.

Going forward, it will be essential to foster interdisciplinary collaboration between the NLP research community, the Persian movie industry, and the engaged user base. By bridging this gap and creating synergies, we can unlock the full potential of NLP to support the continued growth and global competitiveness of the Persian-language movie industry.

Declaration

The authors declare they have used AI language models to provide editorial assistance with language clarity in preparing this manuscript.

Author Contributions

Ramin Safa conceptualized the study, supervised the research, and contributed to writing the manuscript. Sedigheh Kaveh, contributed to data collection, analysis, and assisted in manuscript preparation. Both authors approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

The authors thank the editors and anonymous reviewers for their valuable feedback.

References

[1] Dashtipour, K., Gogate, M., Adeel, A., Larijani, H., & Hussain, A. (2021). Sentiment analysis of persian movie reviews using deep learning. *Entropy*, *23*(5), 596. https://doi.org/10.3390/e23050596

[2] Rasool, G., & Pathania, A. (2021). Reading between the lines: untwining online user-generated content using sentiment analysis. *Journal of research in interactive marketing*, *15*(3), 401–418. DOI:10.1108/JRIM-03-2020-0045

[3] Danyal, M. M., Khan, S. S., Khan, M., Ullah, S., Mehmood, F., & Ali, I. (2024). Proposing sentiment analysis model based on BERT and XLNet for movie reviews. *Multimedia tools and applications*, *83*(24), 64315–64339. DOI:10.1007/s11042-024-18156-5

[4] Pang, B., Lee, L., & Vaithyanathan, S. (2002). *Thumbs up? Sentiment classification using machine learning techniques*. ArXiv Preprint Cs/0205070. https://arxiv.org/abs/cs/0205070

[5] Ruelens, A. (2022). Analyzing user-generated content using natural language processing: a case study of public satisfaction with healthcare systems. *Journal of computational social science*, *5*(1), 731–749. DOI:10.1007/s42001-021-00148-2

[6] Hutto, C. J., & Gilbert, E. (2014). *VADER: a parsimonious rule-based model for sentiment analysis of social media text* [presentation]. Proceedings of the 8th international conference on weblogs and social media, ICWSM 2014 (Vol. 8, pp. 216–225). DOI: 10.1609/icwsm.v8i1.14550

[7] Denecke, K. (2008). Using sentiwordnet for multilingual sentiment analysis. *2008 IEEE 24th international conference on data engineering workshop* (pp. 507–512). IEEE. DOI: 10.1109/ICDEW.2008.4498370

[8] Raees, M., & Fazilat, S. (2024). *Lexicon-based sentiment analysis on text polarities with evaluation of classification models*. ArXiv Preprint ArXiv:2409.12840. http://arxiv.org/abs/2409.12840

[9] Aziz, M. M., Purbalaksono, M. D., & Adiwijaya, A. (2023). Method comparison of Naïve Bayes, logistic regression, and svm for analyzing movie reviews. *Building of informatics, technology and science (BITS)*, *4*(4), 1714–1720. DOI:10.47065/bits.v4i4.2644

[10] AlBadani, B., Shi, R., & Dong, J. (2022). A novel machine learning approach for sentiment analysis on twitter incorporating the universal language model fine-tuning and SVM. *Applied system innovation*, *5*(1), 13. DOI:10.3390/asi5010013

[11] Mahadevaswamy, U. B., & Swathi, P. (2022). Sentiment analysis using bidirectional LSTM network. *Procedia computer science*, *218*, 45–56. DOI:10.1016/j.procs.2022.12.400

[12] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: pre-training of deep bidirectional transformers for language understanding. *NAACL hlt 2019 - 2019 conference of the north american chapter of the association for computational linguistics: human language technologies - proceedings of the conference* (Vol. 1, pp. 4171–4186). Minneapolis, Minnesota. https://eva.fing.edu.uy/pluginfile.php/524749/mod_folder/content/0/BERT Pre-training of Deep Bidirectional Transformers for Language Understanding.pdf

[13] Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2017). *A practical guide to sentiment analysis* (Vol. 5). Springer.

[14] Alhamed, K. (2015). *Framework for open-source software update mechanism : resistance to stacking the deck attack*. Florida Institute of Technology. https://www.proquest.com/openview/3c2c67a5c27a216bee01e0dd82b5816d/1?pq origsite=gscholar&cbl=18750

[15] Sharma, N. A., Ali, A. B. M. S., & Kabir, M. A. (2024). A review of sentiment analysis: tasks, applications, and deep learning techniques. *International journal of data science and analytics*, 1–38. DOI:10.1007/s41060-024-00594-x

[16] Dongo, I., Cardinale, Y., Aguilera, A., Martinez, F., Quintero, Y., Robayo, G., & Cabeza, D. (2021). A qualitative and quantitative comparison between Web scraping and API methods for Twitter credibility analysis. *International journal of web information systems*, *17*(6), 580–606. DOI:10.1108/IJWIS-03-2021-0037

[17] Zablith, F., & Osman, I. H. (2019). ReviewModus: text classification and sentiment prediction of unstructured reviews using a hybrid combination of machine learning and evaluation models. *Applied mathematical modelling*, *71*, 569–583. DOI:10.1016/j.apm.2019.02.032

[18] Zhang, M. L., Peña, J. M., & Robles, V. (2009). Feature selection for multi-label naive Bayes classification. *Information sciences*, *179*(19), 3218–3229. DOI:10.1016/j.ins.2009.06.010

[19] Stiglic, G., Kocbek, S., Pernek, I., & Kokol, P. (2012). Comprehensive decision tree models in bioinformatics. *PLoS one*, *7*(3), e33812. DOI:10.1371/journal.pone.0033812

[20] Sutskever, I. (2013). *Training recurrent neural networks*. University of Toronto Toronto, ON, Canada. https://www.cs.utoronto.ca/~ilya/pubs/ilya_sutskever_phd_thesis.pdf

[21] Fernando, T., Denman, S., McFadyen, A., Sridharan, S., & Fookes, C. (2018). Tree memory networks for modelling long-term temporal dependencies. *Neurocomputing*, *304*, 64–81. DOI:10.1016/j.neucom.2018.03.040

[22] Tabinda Kokab, S., Asghar, S., & Naz, S. (2022). Transformer-based deep learning models for the sentiment analysis of social media data. *Array*, *14*, 100157. DOI:10.1016/j.array.2022.100157

[23] Taboada, M., Brooke, J., & Voll, K. (2011). Lexicon-based methods for sentiment analysis. *Computional linguistics*, *37*(2), 267–307. https://www.proquest.com/docview/896181231/C4F09CD9F4A6440APQ/13?accountid=13827

[24] Zimbra, D., Abbasi, A., Zeng, D., & Chen, H. (2018). The state-of-the-art in twitter sentiment analysis: a review and benchmark evaluation. *ACM transactions on management information systems*, *9*(2), 1–29. DOI:10.1145/3185045

[25] Brzezinski, D., & Stefanowski, J. (2017). Prequential AUC: properties of the area under the ROC curve for data streams with concept drift. *Knowledge and information systems*, *52*(2), 531–562. DOI:10.1007/s10115-017-1022-8

[26] Ariai, F., Mahmoudi, M. T., & Moeini, A. (2024). Enhancing aspect-based sentiment analysis with ParsBERT in Persian language. *Journal of artificial intelligence and data mining*, *12*(1), 1–14. DOI: 10.22044/jadm.2023.13666.2482

[27] Mohtaj, S., Roshanfekr, B., Zafarian, A., & Asghari, H. (2018). Parsivar: a language processing toolkit for persian. *Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018)*. European Language Resources Association (ELRA). https://aclanthology.org/L18-1179/

[28] Oji, R., Razavi, S. F., Dehsorkh, S. A., Hariri, A., Asheri, H., & Hosseini, R. (2021). Parsinorm: a persian toolkit for speech processing normalization. *2021 7th international conference on signal processing and intelligent systems (ICSPIS)* (pp. 1–5). IEEE. DOI: 10.1109/ICSPIS54653.2021.9729392

[29] Youvan, D. C. (2024). Understanding sentiment analysis with VADER: a comprehensive overview and application. https://www.researchgate.net/publication/381650914

[30] Papia, S. K., Khan, M. A., Habib, T., Rahman, M., & Islam, M. N. (2024). DistilRoBiLSTMFuse: an efficient hybrid deep learning approach for sentiment analysis. *PeerJ computer science*, *10*, e2349. DOI:10.7717/peerj-cs.2349

[31] Dadlani, A., Vo, V., Khemka, A., Harvey, S. T., Kyzy, A. K., Jones, P., & Verhoeven, D. (2024). *Leading by the nodes: a survey of film industry network analysis and datasets*. ArXiv Preprint ArXiv:2408.06393. https://arxiv.org/abs/2408.06393